Integrating ToxCast Assays into an Androgen Receptor (AR) Pathway Model

N Kleinstreuer[1], D Filer[2], M Martin[2], K Houck[2], D Zang[1], D Allen[1], W Casey[3], R Judson[2]

[1]ILS/NICEATM, RTP, NC, USA; [2]EPA/ORD/NCCT, RTP, NC, USA; [3]NIH/NIEHS/DNTP/NICEATM, RTP, NC, USA

Abstract

The Tox21 and ToxCast programs include multiple *in vitro* assays conducted in a high-throughput screening (HTS) format that are relevant to the AR pathway and can be used to identify substances with potential androgenic/anti-androgenic activity *in vivo*. Here we used a number of assays that map to the androgen receptor (AR) pathway to build a mathematical model that attempts to distinguish true AR pathway activity from technology-specific assay interference. This battery of nine assays (five from ToxCast and four from Tox21) probes perturbations of the AR pathway at multiple points (receptor binding, cofactor recruitment, gene transcription and protein production) in multiple cell types. We compiled a list of putative AR reference chemicals from the ICCVAM (2003) and OECD (2010) reference chemical lists that includes agonists, antagonists, selective androgen receptor modulators (SARMs), and inactive chemicals. The model showed 96% (23/24) concordance across the reference set, including successfully identifying multiple SARMs with both agonist and antagonist activity. However, fluoranthene, a SARM, was active only in the cofactor recruitment assays and was therefore mispredicted by the model as acting via an assay-specific interference pathway. All chemicals in the ToxCast library known to target AR were correctly identified by the model. We will discuss a variety of patterns of assay activity and pathway predictions across 1846 ToxCast chemicals, and identify those prioritized to be active against the AR pathway. Where available, we will compare predictions to toxicity data from the literature and look for potential trends relating to use case and exposure scenarios.

Introduction

- U.S. (7 U.S.C. 136, 110 Stat 1613) and international regulations require the testing of certain chemicals for the detection of potential endocrine activity (estrogen, androgen, steroidogenesis, and thyroid pathways).
- As many as 10,000 chemicals may lack sufficient testing data, with several hundred new chemicals being added each year (EPA 2011).
- The EPA National Center for Computational Toxicology (NCCT) and the NIH National Center for Advancing Translational Science (NCATS) run multiple endocrine related high throughput screening (HTS) assays as part of the ToxCast and Tox21 research programs.
- Following the estrogen receptor pathway model approach (Judson et al. manuscript in preparation), we have constructed a mathematical model to predict chemical-induced androgen receptor (AR) activity based on nine HTS assays that map to the AR pathway.

Data Sources

- The data used were generated by the U.S. EPA ToxCast chemical research program (Dix et al. 2007; Judson et al. 2010) and the Tox21 federal partnership (Tice et al. 2013).
- Concentration-response data on 1846 chemicals were generated with each chemical tested in up to nine AR pathway assays. Assay technologies included:
  - Two cell-free biochemical radioligand AR binding assays (Novascreen: Knudsen et al. 2011; Sipes et al. 2013)
  - Two cofactor recruitment assays that measure protein: protein interaction between AR and SRC1 (Odyssey Thera: Filer et al. manuscript in preparation)
  - One transactivation assay measuring reporter gene levels (Attagene: Martin et al. 2010; Franzosa et al. manuscript in preparation)
  - Two transactivation assays measuring reporter protein level readouts (Tox21: Huang et al. manuscript in preparation)
  - Two transactivation antagonist assays (Tox21: Huang et al. manuscript in preparation)
- The chemicals were run in concentration-response format in all assays except for the cell-free binding assays. These were initially run at a single concentration (25 µM), and if significant activity was seen, the chemical was then run in concentration-response mode.

AR Pathway Assays

- A summary of the *in vitro* AR assays is shown in **Table 1**. Identifiers (ID) map to the model in **Figure 1**.
- All concentration-response assay data were analyzed using the ToxCast data analysis pipeline, which automates the processes of baseline correction, normalization, curve-fitting, and hit-calling, as well as detection of a variety of potential confounders annotated as "caution flags". This pipeline and all raw and processed data and annotations are publically available (http://actor.epa.gov/).

Table 1. Assays Used in the AR Pathway Model

ID	Assay Name	Source	Gene	Species	Type
A1	NVS human AR	Novascreen	AR	Homo sapiens	Receptor Binding
A2	NVS chimpanzee AR	Novascreen	AR	P. troglodytes	Receptor Binding
A3	OT_AR_ARSRC1_0480	Odyssey Thera	AR;SRC	Homo sapiens	Cofactor Recruitment
A4	OT_AR_ARSRC1_0960	Odyssey Thera	AR;SRC	Homo sapiens	Cofactor Recruitment
A5	ATG_AR_TRANS	Attagene	AR	Homo sapiens	RNA Reporter Gene
A6	Tox21_AR_BLA_Agonist_ratio	NCGC	AR	Homo sapiens	β-Lactamase Reporter Gene
A7	Tox21_AR_LUC_MDAKB2_Agonist	NCGC	AR	Homo sapiens	Luciferase Reporter Gene
A8	Tox21_AR_BLA_Antagonist_ratio	NCGC	AR	Homo sapiens	β-Lactamase Reporter Gene
A9	Tox21_AR_LUC_MDAKB2_Antagonist	NCGC	AR	Homo sapiens	Luciferase Reporter Gene

Cytotoxicity Filter

- For many chemicals, there are many assay hits for both AR and non-AR assays in the concentration range in which cytotoxicity is observed.
- We have developed a scheme to filter out these nonselective assay hits using the mean logAC50(cytotox), the median absolute deviation (MAD) of the logAC50(cytotox) hits, and the median of the MAD of the logAC50(cytotox) distributions across all chemicals (the global cytotoxicity MAD).
- For chemicals with two or more positive responses in cytotoxicity assays, we calculate a "Z-score" for each AR pathway assay hit:

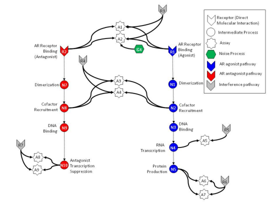
$$Z(chemical, assay) = \frac{\log AC50(chemical, assay) - median[\log AC50(chemical, cytotox)]}{global\,cytotoxicity MAD}$$

- A hit with a large Z-value occurs at concentrations significantly below where cytotoxicity is occurring. This hit is both unlikely to be caused by cell-stress or cytotoxicity-related processes and is more likely to cause toxicity through a target-selective mechanism.

AR Pathway Network

- The graphical representation of the network used to evaluate the integrated *in vitro* assay responses is shown in **Figure 1**. The model was based on the series of molecular events that typically occur in a receptor-mediated response.
  - The process starts with the interaction of a chemical with a nuclear AR (receptor node R1).
  - For example, an AR agonist will cause the receptors to dimerize (node N1), translocate to the nucleus and recruit co-factors to form the complete active transcription factor complex (TF) (node N2).
  - This TF then binds to the chromatin DNA (node N3), initiates transcription of mRNA (node N4) and subsequent translation to protein (node N5).
- Each of these processes (with the exception of dimerization and DNA binding) was measured in the current collection of nine *in vitro* assays represented as white stars.
- The AR pathway is shown in two modes: agonist (blue, acting through R1) and antagonist (red, acting through R2). The model assumes that a chemical that interacts with the AR will bind in one or both of the agonist or antagonist conformations and that this will trigger activity in the appropriate pathway.
- Every *in vitro* assay is subject to processes that can lead to nonspecific activity, independent of the AR pathway node that it is supposed to measure. The assay interference pathways were modeled as alternate "pseudo-receptors" (gray arrow nodes).
- Every *in vitro* assay is also subject to artifacts and sources of experimental noise, and these noise processes are represented by the green hexagon.

Figure 1. AR Pathway Model



Colored arrow nodes represent "receptors" with which a chemical can directly interact. Colored circles represent intermediate biological processes that are not directly observable. White stars represent the assays that measure activity at the biological nodes. Arrows represent transfer of information. The green hexagon represents a noise process to which the assays are subject. Only a single example is explicitly shown, but each assay has its own underlying noise process.

Mathematical Model

- The computational model assumes that the value (the efficacy, *A*) returned by an assay at a given concentration is a linear sum of the contributions from the receptors that it measures (i.e. it is a simple linear additive model):

$$A_i = \sum_j F_{ij} R_j$$

- The goal is then to find a set of values that minimize the difference between the predicted assay values (A_i^{pred}) and the measured ones (A_i^{meas}) for each chemical and concentration. For each chemical-concentration pair, a constrained least-squares minimization approach is used, where the function being minimized is:

$$\varepsilon^2 = \sum_i \left(A_i^{pred} - A_i^{meas}\right)^2 + penalty(R)$$

- The term *penalty(R)* penalizes solutions that predict that many receptors are being simultaneously activated by the chemical. It is given by:

$$penalty(R) = \alpha \frac{(SR^2)}{(SR^2 + SR_0^{\,2})}$$

- In this equation, *SR* is the sum of R values at that concentration, SR_0 is a threshold value and α is a small number between 0 and 1. This penalty term helps stabilize the solutions and enforces a reasonable physical assumption about chemical promiscuity, i.e., that it is unlikely that most chemicals will strongly and specifically interact with many dissimilar molecular targets.
- The model results in a response value (between 0 and 1) for each receptor at each concentration. These results are summarized as area under the curve (AUC), which is the integral across the concentration range:

$$AUC_j = \frac{1}{N_{conc}} \sum_{i=1}^{N_{conc}} sign(slope) \times R_j(conc_i)$$

Reference Chemical Performance

- A set of 24 positive and negative reference chemicals were used to evaluate the performance of the model.
- These reference chemicals were identified based on reports from ICCVAM (ICCVAM 2003) and OECD (OECD 2010). Chemicals were chosen that had consistent *in vitro* results across both reports and that were also in the ToxCast library.
- The reference chemicals and their predicted androgen agonist and antagonist activities are given in **Table 2**.

Table 2. Reference Chemicals

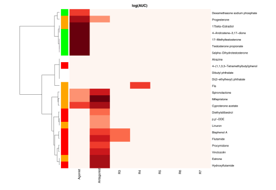
CAS RN	Chemical Name	Activity
2392-39-4	Dexamethasone	Agonist
63-05-8	4-Androstenedione	Agonist
521-18-6	5a-Dihydrotestosterone	Agonist
58-18-4	Methyl testosterone	Agonist
57-85-2	Testosterone propionate	Agonist
13311-84-7	Flutamide	Antagonist
140-66-9	4-tert-Octylphenol	Antagonist
32809-16-8	Procymidone	Antagonist
80-05-7	Bisphenol A	Antagonist
50471-44-8	Vinclozolin	Antagonist
72-55-9	p,p'-DDE	Antagonist
52806-53-8	Hydroxyflutamide	Antagonist
56-53-1	Diethylstilbestrol	Antagonist
84-74-2	Di-n-butyl-phthalate	Inactive
117-81-7	Diethylhexyl phthalate	Inactive
1912-24-9	Atrazine	Inactive
427-51-0	Cyproterone acetate	SARM
50-28-2	17-β-Estradiol	SARM
53-16-7	Estrone	SARM
330-55-2	Linuron	SARM
52-01-7	Spironolactone	SARM
57-83-0	Progesterone	SARM
84371-65-3	Mifepristone	SARM
206-44-0	Fluoranthene	SARM

Abbreviations: CAS RN = Chemical Abstracts Service Registry Number; SARM = selective androgen receptor modulator, which has both agonist and antagonist activity.

Model Results

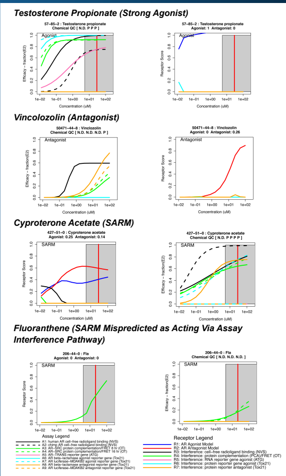
- The AR pathway model predictions are shown in **Figure 2** as a heatmap. The chemicals are plotted against their receptor AUC values, with R1 being agonism and R2 being antagonism.
- Overall, the model showed 96% (23/24) concordance in identifying agonist or antagonist AR activity across the reference set, using a threshold of 0.01 as a positive AUC score.
  - The three inactive reference chemicals were identified by the model as being inactive.
  - All five agonist reference chemicals produced a high R1 score, and did not show any patterns of assay interference.
  - Of the eight antagonist reference chemicals, all were identified as antagonists with R2 scores greater than 0.01. In **Figure 2**, it appears that 4-(1,1,3,3-tetramethylbutyl)phenol (4-tert-octylphenol) was inactive but that is due to a threshold issue where only scores >0.05 were plotted; this chemical has an antagonist model score of R2 = 0.036.
  - Two antagonist reference chemicals, bisphenol A and flutamide, were also predicted to potentially act via assay interference pathways, but the R3 model scores were lower than for R2 (antagonism).
- The model successfully identified multiple selective androgen receptor modulators (SARMs) with both agonist and antagonist activity.
  - Four SARMs were correctly predicted to have both agonist and antagonist activity by the model, while two SARMs (estrone and linuron) were only identified as antagonists and one SARM (17-β-estradiol) was only predicted to be an agonist.
  - Fluoranthene, also a SARM, was active in the cofactor recruitment assays but none of the other AR pathway assays and was therefore mispredicted by the model as acting via an assay-specific interference pathway.
- Examples of assay concentration-response plots and model AUC predictions are shown in **Figure 3** for testosterone propionate (agonist), vincolozolin (antagonist), cyproterone acetate (SARM), and fluoranthene (SARM, missed by the model).

Figure 2. AR Pathway Receptor AUC Values for Reference Chemicals



Reference chemicals are color-coded on the sidebar as agonist (green), antagonist (red), SARM (both agonist and antagonist activity, orange) or inactive (white). In the heatmap, darker red indicates larger AUC values. A minimum cutoff of 0.05 AUC score was used to generate the heatmap.

Figure 3. Examples of Reference Chemical Activity in Assays and Receptor AUC Values From the AR Pathway Model



Acknowledgements

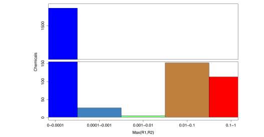
The Intramural Research Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by ILS under NIEHS contracts N01-ES 35504 and HHSN27320140003C.

The views expressed above do not necessarily represent the official positions of any Federal agency. Since the poster was written as part of the official duties of the authors, it can be freely copied.

AR Pathway Activity Across the ToxCast Library

- **Figure 4** shows the distribution of the AR model pathway scores (the maximum agonist or antagonist score for each chemical) across the ToxCast chemical library.
- Of the 1846 chemicals tested in the AR pathway assays, 1549 were inactive in the model, with both R1 and R2 scores below 0.0001, while 115 chemicals were predicted to strongly affect the pathway either as agonists or antagonists (R1 or R2 >0.1). The remaining 182 chemicals had model scores in the intermediate region.

Figure 4. AR Pathway Model Scores for 1846 ToxCast Chemicals



The histogram shows AR pathway model scores, using the maximum R1 (agonist) or R2 (antagonist) value and without applying the cytotoxicity filter, across the 1846 chemicals in the ToxCast library.

Conclusions

- The AR pathway model performed well against the reference chemical set, including identifying SARMs with both agonist and antagonist activities. Further, all 15 compounds in the library whose target gene is known to be AR were identified by the model as either agonists or antagonists with R1 or R2 >0.05.
- The majority of ToxCast chemicals tested in the AR assays were predicted to be inactive against the pathway. Certain environmental chemicals such as antimicrobials (e.g. triclosan and triclocarban) and plasticizers (e.g. bisphenol A and bisphenol AF) were predicted to be AR antagonists; however, this was confounded by cytotoxicity and may require more targeted testing within the relevant concentration ranges.
- The AR pathway model provides a biologically-based mathematical approach to distinguish assay interference from true agonist or antagonist activity, and to prioritize large numbers of environmental chemicals for their potential androgenic or anti-androgenic activity.

References

Dix et al. 2007. Toxicol Sci 95(1): 5-12.

Food Quality Protection Act of 1996. 7 U.S.C. 136. Public Law 104-170.

EPA. 2011. The Incorporation of *In Silico* Models and *In Vitro* High Throughput Assays in the Endocrine Disruptor Screening Program for Prioritization and Screening. Available: http://www.epa.gov/endo/pubs/edsp21_work_plan_summary%20_overview_final.pdf.

ICCVAM. 2003. ICCVAM Evaluation of *In Vitro* Test Methods for Detecting Potential Endocrine Disruptors. National Institute of Environmental Health Sciences.

Judson et al. 2010. Environ Health Perspect 118(4): 485-492.

Knudsen et al. 2011. Toxicology 282(1-2): 1-15.

Martin et al. 2010. Chem Res Toxicol 23(3): 578-590.

OECD. 2010. Draft Report of Pre-validation and Inter-laboratory Validation for Androgen Receptor (AR) Mediated Stably Transfected Transcriptional Activation (AR-STTA) Assay to Detect Androgenic and Anti-Androgenic Activities. OECD Publishing.

Safe Water Drinking Act Amendments of 1996. 110 Stat. 1613. Public Law 104-182.

Sipes et al. 2013. Chem Res Toxicol 26(6): 878-895.

Tice et al. 2013. Environ Health Perspect 121(7): 756-765.